A Commercial Implementation of the JPL Virtual Reality Calibration Technology for Telerobotics

Won S. Kim
Jet Propulsion Laboratory
M/S 198-219
4800 Oak Grove Drive
Pasadena, California 91109
email: wonsoo@telerobotics.jpl.nasa.gov

Robert Brown, Brian Christensen, Chris Beale
Deneb Robotics, Inc.
3285 Lapeer Road West
P. O. Box 214687
Auburn Hills, Michigan 48321-4687
email: bob@deneb.com

Abstract

There is a growing interest in the virtual reality (VR) calibration technique of matching graphically simulated virtual environments in 3-D geometry and perspective with actual video camera views. Jet Propulsion laboratory (JPL) recently developed such a technique that enables high-fidelity preview/predictive displays with reliable, accurate calibrated graphics overlay on live video for telerobotic applications, and demonstrated its effectiveness in a recent JPL/NASA-GSFC (Goddard Space Flight Center) remote servicing task. Within NASA's recent thrust for industrial collaboration, JPL recently established a technology cooperation agreement (TCA) with Deneb Robotics, Inc. In this JPL-Industry cooperative Deneb Commercialization Task, JPL transfers the VR calibration software technology to Deneb, and Deneb inserts this software technology into its commercial product. This joint technology collaborative work will enable Deneb to commercialize an upgraded industry product that will greatly benefit both space and terrestrial telerobotic applications. On-going new developments of semi-automatic VR calibration techniques using multi-resolution correlation-based area matching and edge-based feature matching are also presented as evolving technical additions to enhance the existing operator-interactive VR calibration technology significantly.

I. Introduction

Graphic simulation has been widely used in telerobotic applications during the off-line task analysis and planning and also during the introductory operator training. However, the use of graphic simulation during the on-line telerobotic operation, for example, as a tool for on-line preview and predictive visualization, has been limited due to the lack of accurate matching between the simulated environment and the actual remote site task environment, J PL recently developed a virtual reality (VR) calibration technique [16]-[17] that enables reliable and accurate matching by operator-interactive camera calibration and object localization procedures with new linear/nonlinear least-squares algorithms for multiple-camera views. This VR calibration capability enables accurate visual planning, preview, and prediction of robot motion by overlaying virtual graphical images on real images, providing powerful new graphics-based tools for real-time simulation and control of robots in both terrestrial and space applications. A recent JPL/NASA-GSFCORU (Orbital Replacement Unit) changeout remote servicing task performed in May, 1993 demonstrated the usefulness of the developed VR calibration technique and its application to preview/predictive displays with calibrated graphics overlay on live video.

This paper describes our on-going J PL/Deneb joint technology collaborative work that is enabling Deneb to commercialize an upgraded industrial product with a VR calibration video overlay option. The first release of the option, which will be available by the end of February 1995, will be limited to "manual operator-interactive" VR calibration for fixed cameras. The second release, which is planned to be available by September 1995, will include "semi-automatic" VR calibration capability for both fixed and moving cameras as an evolving technical addition to the first year's "manual operator-interactive" VR calibration option. Section 2 briefly summarizes the previously demonstrated JPL VR calibration technique, and Section 3 describes the actual on-going implementation of the VR calibration technique on a Deneb's commercial product TELEGRIP. Section 4 presents our current new developments of semi-automatic VR calibration techniques by using correlation-based area matching and edge-based feature matching. Section 5 illustrates future planned work and potential space and terrestrial applications. The conclusion appears in Section 6.

II. Virtual Reality Calibration

The existing VR calibration technique that was used in tile JPL/NASA-GSFC telerobotic demonstration enables reliable, accurate matching through operator-interactive camera calibration and object localization procedures. These are briefly described **here.**

2.1 Camera Calibration Using a Robot Arm

Our camera calibration method which is designed for calibrated graphics overlay has three key new features; 1) A robot arm itself is used as the calibration fixture, eliminating cumbersome procedures of using an external calibration fixture. 2) An operator-interactive data entry is adopted to obtain reliable correspondence data, since it is still difficult for a computer vision system to find correspondence points reliably. 3) A nonlinear least-squares algorithm combined with a linear least-squares one is employed to obtain accurate camera parameters, where the linear least-squares solution is used as an initial guess. Once the camera parameters are obtained relative to the robot arm base frame through this camera calibration procedure, the graphics model of the robot arm can be overlaid on the video camera view. Details of the operator-interactive linear/nonlinear camera calibration algorithms and their software listings can be found in the recent JPL report [15], which was prepared as part of the JPL-Industry cooperative Deneb Commercialization Task.

2.2 Object Localization

In the original predictive display, only the robot arm graphic model was overlaid on live video as a predictor of the time-delayed robot motion. In our new approach, the object localization procedure has been added after the camera, calibration to determine the object pose (position and orientation) and enable graphic overlay of both the robot arm and the object (s) on live video. Since the object pose becomes known through object localization, our new approach enables the semi-automatic computer-generated trajectory mode in addition to the teleoperation mode. Our object localization method has three key new features; 1) An operator-interactive method is adopted to obtain reliable correspondence data, 2) A projection-based linear least-squares algorithm is extended to handle multiple camera views.

3) A nonlinear least-squares algorithm combined with the extended linear one is employed to obtain an accurate object pose from multiple camera views. Details of these projection-based linear/nonlinear least-squares algorithms and their software listings can be found in [1 5].

The above VR calibration techniques were applied to the JPL/NASA-GSFC remote servicing demonstration. An example of a calibrated video overlay after the VR calibration is shown in Fig. 1, where both the robot arm ancl the object graphic models arc superimposed on the video image. Experimental measurements of calibration errors in inserting a tool into the ORU hole indicated that the positioning alignment accuracy achieved by the developed calibration technique using four camera views was 5.1 *mm* on the average with a 10.7 *mm* maximum error at 95% confidence level. The depth error was 6.5 *mm* on the average, with a 13.7 *mm* maximum depth error at 95% confidence level.

2.3 **Preview/Predictive Displays**

After matching 3-1) graphic models of a virtual environment with actual camera views through the above VR calibration technique, the operator can now perfo m a telerobotics servicing task by using preview/predictive displays with calibrated graphics overlay on live video. l'review/predictive displays allow the operator to generate the simulated robot arm trajectory in preview and then to visually monitor and verify the actual remote robot arm motion with confidence, thus enhancing safety and reliability in remote servicing operations regardless of communication time delay. Our preview/predictive display is useful not only for non- contact tasks but also for contact or insertion tasks involving compliante/impedance control in the remote site. This is because the simulated graphics arm is updated with the actual final robot joint angle values after the completion of each robot arm trajectory command at the remote site. This update eliminates accumulation of small motion execution errors as well as large compensation errors due to the compliance/impedance control. Fig. 2 shows an example of a preview/predictive display during the performance of the JPL/NASA-GSFC demonstration task.

III. Irnplementation on TELEGRIP

3.1 Open Architecture

The TELEGRIP Access framework [3] is an open architecture based upon Dynamic Shared Objects (11 S0's) as shown in Fig. 3. 1)S0's provide many benefits when compared with other strategies for incorporating user-defined modules with a centralized kernel including speed of development, access to all internal functions and data including the entire geometric database, flexibility in development, and minimizing platform dependence. The use of DSO's inc eases speed of development because the developer does not have to link a fixed library every time a change is desired. The mapping of user libraries to application libraries occurs at run time. This enables the developer to compile and load only the code that has changed. The use of DSO libraries gives the Access developer the same level of system functionality that original developers have. Internal functions and data are accessed directly with no degradation in performance.

A key feature provided by the TELEGRIP open architecture is that it allows developers/users to add their own virtual reality calibration algorithms and video overlay methods, if necessary.

3.2 Camera Viewing Model

The TELEGRIP viewing model [4] for perspective projection is defined by a pyramid which is inscribed within the viewing cone (Fig. 4). Only the geometry inside the pyramid truncated by hither and yonder planes (viewing frustum) is visible. The user can specify the perspective projection parameters such as image plane size (IIPS), focal length (II'],), and field of view angle (FOVY), which are related by FOVY = 2 * atan(IIPS/(2*FI)).

3.3 Muti-Window Video Overlay

The TELEGRIP Access API (application programmers interface) provides a rich suite of functionality which can be used to replace and enhance internal functions, access and extend the internal database, and rapidly develop toolkit applications, all with the guarantee of portability between TELEGRIP releases. Video overlay leverages upon Access to create and manipulate multiple graphics windows, en-

hance rendering functions to effect video/graphics blend, and extend the internal database to maintain video calibration data within the TELEGRIP camera viewing data strut.turcs.

The video overlay procedure starts with the definition phase in which a standard "User View", TELEGRIP's representation of a virtual camera, is created. This can be accomplished either programmaticall y through Access function calls, or through the traditional TELEG R]}'button interface. The user then defines a live video or a captured 'video image source for the view which is stored within the view data structure as an Access database extension. The view may then be displayed in an external window, either in graphics, video, or blended graphics/video display mode. Blending is accomplished via a Video Overlay function registered as an Access pre-display callback system which draws the video into the frame buffer just before graphics scene is rendered. When the operator wishes to proceed with the calibration phase of video overlay, the view window to be calibrated is automatically placed into video display mode. The operator then chooses 3D graphics points from any window currently in graphics display mode and correlates them with 2D video points from the calibration view window. This data is stored with the video setup in view structure Access database extension. Upon completion of data point acquisition, the view is calibrated and placed into blended ,g]al)hics/video display mode from which the operator may inspect the calibration results.

The TELEGRIP video overlay implementation is based upon an application programmers interface (API) layer which insulates the overlay developer from the specifics of video hardware, thus enabling support over a wide range of video products. Support is currently planned for the SGI VideoLab, Galileo, Indigo2, lnd y, and Serius Videoboards encompassing t he entire range of current SGI computing hardware from the Indy to the Onyx.

IV. Semi-Automatic Virtual Reality Calibration

Semi-automatic and automatic techniques of matching graphic model images (virtual environments) to actual video images by using model-based image processing can significantly enhance the current "opcra,tor-interactive" VR calibration technique for telerobotic applications. Two primary image matching methods are 1) area matching that uses correlations between image regions, and 2) feature matching

that mat ches features such as edges and corner points between images. 1 n general, area-based and feature-based matching arc considered to be complementary rather than competing with each other. Both area and feature matching methods have been implemented recently to compare and take advantage of both methods. Preliminary results are included in this paper, but further tests and improvements are necessary.

4.1 Correlation-Based Area Matching

The correlation-based area matching algorithm that we implemented employs three key techniques: 1) multi-resolution hierarchical coarse-to-fine strategy, 2) edge-based block matching, and 3) principal axes procedure using eigen vectors of each correlation matrix. The multi-resolution coarse-to-fine strategy [1], [8], [20] provides two advantages; 1) coarse global matching using lower resolution images with a larger size of the search area guides fine local matching using higher resolution images with a smaller size of the search area, reducing the number of false matches, and 2) the correlation-based matching speed increases markedly. Block matching [21] has been extensively used in today's video compression standards such as MPEG. In block matching, a reference image is segmented into small rectangular blocks, and for each block one displacement vect or is calculated that represents the shift of the image block to match best in the second image. Recently edge-based block matching, instead of the conventional intensity-based block matching, has been successfully used for computational efficiency with similar matching performance characteristics [25]. Correlation matching of image areas containing edges yields the displacement vector that has a high confidence of accuracy along the edge direction but a very low confidence along the perpendicular direction. These different levels of confidence in accuracy can be incorporated into the correlation matching algorithm by placing different weights, depending upon their eigen values, along the principal and minor eigen vector axes of the correlation matrix [24]. The details of the algorithm implemented are described here.

1, Construct multi-resolution images from full-resolution images (640 pixels x 480 pixels) - 320x240 (1/2 scale for both x and y axes), 160x120(1/4 scale), 80x60 (1/8 scale), and 40x30 (1 /16 scale) images for the graphics model image and also for the video image. Apply a,) Sobel or b) Canny edge detector to the graphic. model images of different resolutions and generate binary edge-detected

graphic model images. Apply a) Sobel or b) Canny gradient operator to the video images of different resolutions and generate gray-scale gradient video images.

2. Start the correlation matching from the I/S-scale resolution images (the l/16-scale images were not used due to poor correlation results). Form a sampling grid with a sampling interval of 5 pixels in both column and row axes of the graphic model image for matching. The template image size is selected to be a 9x9 block.

3. Start the correlation matching from the 1/8-scale images (1/1 6-scale images were not used due to poor correlation results). Segment the graphic model image into overlapping 9x9 windows as template images with an inter-window distance of 5 pixels both horizontally and vertically. For each template image that has at least one edge pixel, obtain the 11x1 1 correlation matrix by computing the correlation coefficients between the templat e image displaced and the corresponding area in the video image, where the search area window size in the video image is selected to be 19x1 9, and thus the displacement of the 9x9 template image is up to ± 5 pixels both horizontally and vertically. At present, the mean and variance normalized correlation [2], [10] is used as a correlation measure. If the maximum correlation coefficien t of the correlat ion matrix is greater than 0.5, the template image is considered] to have the best match at the corresponding displacement, which is represented by image displacement vector.

4. Repeat Steps 2 and 3 in higher resolution images to the full-scale highest resolution level. Usc the image displacement vector obtained from the immediate lower resolution to specify the location of the search area for the next higher resolution image.

5. For each template image that has its image displacement vector (with its maximum correlation coefficient greater than 0.5), obtain the 3-D position of an object edge point in the temple image by using the z-buffer (depth) data of the graphic model image. Also compute the corresponding 2-D image point from the image displacement vector.

6. For each template image that has its image displacement vector, compute the major (principal) and minor eigen vectors of its correlation matrix. 'l'he eigen vector associated with the maximum

eigen value indicates the direction of the dominant linea r edge in the match area.

7. Perform the nonlinear least squares object localization using the 3-D object points and 2-D image points. Place lower weights along the principal axes and higher weights along the minor axes.

8. Remove "outliers" to eliminate false matches [7], and then re-do the non-linear least-squares object localization of Step 7.

As an initial test, the above multi-resolution correlation-based matching algorithm has been applied to a calibrated video image used during the performance of the JPL/NASA-GSFC remote servicing demonstration task. A graphics model of an ORU (Orbital Replacement Unit) and a video image of the NASA-GSFC remote site arc shown in Fig. 5a and 5b respectively. The video image contains an ORU, an Explore spacecraft mockup, and an RRC (Robotics Research Corp.) arm with a servicing tool mounted at the end of the arm. 'l'he pose of the ORU graphic model was calibrated originally during the demonstration to match with the video image by using the existing "operator-interactive" VR calibration technique. in Fig. 5a, this calibrated graphic model of" the ORU was intentionally translated by 2 cm along all three x, y, and z translational axes and rotated by 3 degrees about all three x, y, and z rotational axes to test the above correlation-based matching algorithm.

Fig. 6a shows 1 /2, 1 /4, 1/8, and 1 /16 scale graphics model and video images after processing the images with the Canny edge detector. Correlation results obtained from these images are graphically shown in Fig. 6b. Each square represents the 5x5 center portion of the 9x9 template image for which the image displacement vector is found, and the line segment starting from the center of the square represents the displacement vector. Although we can observe several instances of false matches and missing matches in Fig. 6b, Step 8 of the above algor ithm yielded the desired correct solution for the ORU object pose in this example tested for both Sobel and Canny operators, However, other preliminary tests with larger orientation errors indicated that the above cor relation-based matching technique using edge-detected images is very sensitive to the orientation difference between the initial graphic model pose and the actual object pose in the video image, and appears to be useful only for a small initial orientation difference of less than 2 or 3 degrees. For now, this technique appears to be useful only for the fine alignment/matching of the graphic model to the video image. This technique is of course very

useful for object tracking with small displacements between two consecutive video image frames.

4.2 Edge-Based Feature Matching

Feature-based matching of a video image of an object to its geometric model has been widely investigated. In the feature-basul matching, features such as points and edges are used in finding the best match between the image and model features to determine the object pose (position and orientation) or the camera viewing parameters.

For m model features and i image features, the search for the best match in general results in an exponential search problem since some model features may not have corresponding image features and vice versa due to the occlusion by other objects and the noisy edge detection. To cope with this exhaustive search problem, several researchers [6], [11], [13] proposed an efficient search strategy, so called "hypothesize-and-test" strategy, by transforming the exponential search to a polynomial one. The '(hypothesize-ant]-test" strategy is an iterative two-stage search consisting of hypothesis generation and hypothesis test (verification). In the hypothesis generation stage, a new combination of the minimal number of model-image feature pairs is selected to determine the geometric transformation between the object model and its image. The transformation computed is then used in the hypothesis test stage to project the object model features onto the image and find compatible or aligned image features. The model-image alignment is scored by comparing the transformed model features and image features. The best alignment is the one that maps the most model features onto image features. This two-stage procedure is repeated to find a satisfactory match.

Since edges are easier to detect and more reliable than corner points, we use edges for matching. In the edge-basecl feature matching, three pairs of model and image edges are tile minimum number of edge pairs to estimate the geometric transformation [5], [12]. Therefore, the search space is now reduced to find the best triplets of model and image edge pairs, resulting, in a polynomial search. The edge-based feature matching algorithm implemented for our semi-automatic VR calibration is as follows:

1 Obtain two separate lists of edge features, one from the graphic model image and the other from the video image, by detecting edge pixels with a Canny operator, linking neighboring edge pixels, and breaking the chains into approximating straight line segments. The Vista software distribution

package [1 9], [23], which wc obtained over the computer Internet, provides all these capabilities. Its related C routines were used in our application.

2. Sort each edge list in the order of edge length. Remove short edges, for instance, edges smaller than 15 pixels. For the graphic model, merge edges that lie on a straight line with a relatively small broken gap, for instance, less than a 50-pixel gap between the two edges. For the video image, edges on a straight line are not merged, but are considered later in Step 5.

3. Make a search list by finding compatible edges for every model edge. The compatibility criteria used are: 1) orientation difference between the model and image edges less than 20 degrees, and 2) the distance between the model and image edges less than 30 pixels. A simple distance measure from the image edge to the corresponding model edge is the normal distance from the mid point of the image edge to the model edge. Use this measure as the distance between the image and model edges, if the normal projections of the both endpoints of the image edge to the model edge lie inside the model edge segment. If one of the normal projections of the image edge endpoints lie outside the model edge segment, the distance is increased to take into account this offset and is defined as the distance between the protruding image edge endpoint and its associated nearby model image endpoint. If both of the normal projections of the image edge endpoints lie outside the model edge segment, the larger distance of the two image edge endpoints to their associated nearby model edge endpoints is defined as the distance between the image and model edges. This distance computation is based on the fact that normally the image edge will not be longer than the corresponding model edge correctly projected [9].

4. Make a reduced search list from the full search list of Step 3 by imposing further constraints: 1) remove the image edge from the list if the image edge length is greater than 1/4 of or less than 4 times the model edge length, and 2) keep only the longest edge in the list for image edges lying on a straight line. once the full and reduced search lists are found, perform the following iterative hypothesize-and-test procedure to obtain the best match and determine the object pose (or camera viewing parameters). The reduced search list is used in the hypothesis generation stage to significantly reduce the initial search space. The full search list is used during the hypothesis

test stage.

5. Hypothesis generation. Select a triplet of model and image edge pairs in the reduced search list. Drop the triplet if all three model edges are parallel. Also drop the triplet if two of the model edges are parallel and very closely located relative to the model size. In selecting a combination of triplets, NIL or no-match condition must be included to consider occlusions and noisy edge detections.

6. Hypothesis test. For a selected triplet, compute the geometric transformation between the object model and its image. In our approach, we assume that an approximate object pose is (or approximate camera viewing parameters are) known, and a simple non-linear least-squares method is used to compute the rotation first and then the translation. With the obtained transformation, project the object model edges onto the image and find the compatible, aligned image edges that are closest in distance as defined in Step 3. If the closest aligned image edge has other image edges on a straight line, these edges are also considered as aligned edges if they are compatible with the corresponding model edge. The model-image alignment score is the sum of all the aligned edge lengths.

7. Repeat the above hypothesize-and-test procedure of Steps 4 and 5 for the search depth levels of 3, 4, and 5. The minimum search depth of level 3 considers the three longest model edges, where each of the model edges has at least one compatible image edge. To enhance the matching reliability, the search level is expanded to level 5. The best match is the one that produces the highest alignment score.

The above edge-based feature matching is being tested, and a preliminary result is shown in Figs. 7 and 8. Fig. 7 a) and b) show straight line segments detected after Canny and Lowe operators for the graphics model of the ORU and the video image, respectively. The initial pose for the graphics model which was an instance used in our preliminary test was obtained by translating the calibrated graphic model by 5 *cm* along all three x, y, and z translational axes and by rotating the model by -5 degrees about all three x, y, and z rotational axes. Fig. 8a snows a graphic overlay of the graphic model

, at this initial pose on the video image, and Fig. 8b shows a graphic overlay of the pose of the ORU estimated after appying the above edge- based feature matching algorithm. It is significant to note that the algorithm was able to find a good estimate in this example even though about 50% of the visible edges of the ORU model is occluded. The algorthim tends to yield good estimates in most cases tested when the initial pose is within ± 5 *cm* and ± 5 degrees from the actual pose, although in some cases the solution was not the desired one, requiring operator's supervision. Since the area matching and feature matching techniques are complementary, a combination of the correlation and feature matching is anticipated to improve the matching results.

V. Future Work and Potential Applications

Future planned work includes 1) further tests and enhancements of semi-automated VR calibration techniques including object tracking, and 2) on-line interactive model building; and modification.

'l'he results of our joint development efforts will become a key enabling technology in the use and application of augmented reality, In augmented reality implementations, knowledge of a actual environm ent and a virtual world are combined to ease user task execution. We envision wide application of accurate and calibrated virtual worlds to plan and execute complex and dangerous tasks. A flexible and user friendly calibration technique is requisite for this effort. lhc VR calibration video overlay option implemented on TELEGRIP through this collaborative effort will benefit both space and terrestrial telerobotic applications, providing 1) immediate benefits to NASA for ground-controlled telerobotics servicing in space and 2) immediate benefits to the national DOE (Department of Energy) labs working on the disposal and remediation of nuclear waste.

In addition to the above NASA and DOE applications, we see wide applications in execution of otherwise impractical tasks. For example: 1) Casualty Training - the use of a virtual smoke and . fire models to plan and prepare for casu alt y situations in airplanes, boats, tall buildings, 2) Agile Manufacturing -- the use of calibrated virtual worlds to direct manufacturing line reconfiguration with minimal process interruption, 3) Wire Harnesses --- the usc of cabling models to aide operators in the manufacture of complex wire harnesses, 4) Remote Surgery - wc see this technology as key to allowing

manu fact ure of com plex wire harnesses, 4) Remote Surgery — wc see this t echnology as key to allowing remote surgery, transmitted video and accurate computer models will allow safe long distance operation, and 5) Construction - the ability to see where new construction should take place and the ability to track accuracy of construction should be invaluable.

VI. Conclusion

Within NASA's recent thrust for industrial collaboration, JPL and Deneb Robotics, Inc. established a technology cooperation agreement (TCA) on VR (virtual reality) calibration in September 1993. We have taken the following approach in our JPL-Industry cooperative Deneb Commercialization Task: 1) JPL transfers the existing VR calibration technology and its evolving new technical additions to Deneb, 2) Deneb, cooperating with JPL, inserts this software technology into its commercial product TELEGRIP as the VR calibration video overlay option for marketing, and 3) in return, NASA utilizes this enhancement of a commercially supported product for NASA applications, This enhanced commercial product will greatly benefit both space and terrestrial telerobotics. On-going significantly new technical additions include semi-automalic VR calibration techniques using model-based image processing.

Acknowledgment

This work was performed partly at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration, and partly at the Deneb Robotics, Inc., through a JPL-industry Technology Cooperation Agreement (TCA). The authors would like to thank A. K. Bejczy, P. S. Schenker, C. R. Weisbin, and S. K. Khanna of JPL for their advice and encouragement, enabling successful collaborative efforts in the JPL- industry cooperative Deneb Commercialization Task.

References

[1] M. Bierling, "Displacement Estimation by Hierarchical Blockmatching," SPIE vol. 1001 Visual Com-

munications and Image Processing, pp. 942-951, 1988.

[2] 1', J, Burt, C. Yen, and X. Xu, "Local Correlation Measure for Motion Analysis: a Comparative Study," IEEE Conf. on Pattern Recognition and Image Processing, p p. 269-274, 1982.

[3] Deneb Robotics, Inc., *TELEGRIP Access Reference Manual*, 1994.

[4] Deneb Robotics, Inc., *TELEGRIP User Manual and Tutorials*, 1994.

[5] M. Dhome, M. Richetin, J. T. Lapreste, and G. Rives, "Determination of the Attitude of 3-D Objects from a Single Perspective View," IEEE Trans. PAMI, vol. 11, no. 12, PP· 1265-1278, 1989.

[6] 0. D. Faugeras and M. Hebert, "The Representation, Recognition, and Locating of 3-D Objects," Int. J. of Robotics Research, vol. 5, no. 3, pp. 27-35, 1986.

[7] M. A. Fischler and R. C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with applications to image Analysis and Automated Cartography," Comm. ACM, vol. 24, 110. 6, pp. 381-395, 1981.

[8] F. Glazer, G. Reynolds, and P. Anandan, "Scene Matching by Hierarchical Correlation," IEEE Conf. on Computer Vision and Pattern Recognition, PP· 432-441, June 1983.

[9] R. R. Goldberg, "Pose Determination of Parameterized Object Models from a Monocular Image," Image and Vision Computing, 1993.

[10] M. J. Hannah, "A System for Digital Stereo Image hatching,') Photographic Engineering and Remote Sensing, vo. 55, no. 12, pp. 1765-1770, 1989.

[11] R. Horaud, "New Methods for Matching 3-D Objects with Single Perspective Views," IEEE Trans. PAMI, vol. 9, 110.3, pp. 401-412, 1987.

[12] T. S. Huang and A. N. Netravali, "Motion and Structure from Feature Correspondences: A Review," Proc. IEEE, vol. 82, no. 2, pp. 252-268, 1994.

[13] I). P. Huttenlocher and S. Ullman, "Object Recognition Using Alignment," Int. Conf. on Computer Vision, pp. 102-111, London, June 1987.

[14] W. S. Kim, "Virtual Reality Calibration for Telerobotic Servicing," IEEE Int. Conf. on Robotics and Automation, pp. 2769-2775, San Diego, May 1994.

[15] W. S. Kim, *Virtual Reality Calibration: Algorithms and Software Listings with an application to Preview/Predictive Displays for Telerobotic Servicing,* Jet Propulsion Laboratory internal Document 1)-11593, Feb. 1994.

[16] W. S. Kim and A. K. Bejczy, "Demonstration of a High-Fidelity Predictive/Preview Display Technique for Telerobotic Servicing in Space," IEEE Trans. on Robotics and Automation, vol. 9, no. 5, pp. 698-702, 1993.

[17] W. S. Kim, P. S. Schenker, A. K. Bejczy, S. Leake, and S. Ollendorf, "An Advanced Operator Interface Design with Preview/Predictive Displays for Ground-Controlled Space Telerobotic Servicing," SPIE Conference 2057: Telemanipulator Technology and Space Telerobotics, pp. 96-107, Boston, MA, Sept. 1993.

[18] W. S. Kim, H. Seraji, P. Fiorini, R. Brown, B. Christensen, C. Beale, J. Karen, and P. Eismann, "Commercialization of JPL Virtual Reality Calibration and Redundant Manipulator Control Technologies," Third Int. Symp. on Artificial Intelligence, Robotics, and Automation for Space, pp. 23-26, 1994.

[19] D. G. Lowe, *Perceptual Organization and Visual Recognition*, Kluwer, *1985.*

[20] H. P. Moravec, "Towards Automatic Visual Obstacle Avoidance," 5th Int. Joint Conf. on Artificial Intelligence, p. 584, Aug. 1977.

[21] H. G. Musmann, P. Pirsch, and H. Grallert, "Advance: in Picture Coding," Proc. IEEE, vol. 23, no. 4, pp. 523-548, 1985.

[22] N. Navab and O. D. Faugeras, "Monocular Pose Determination from Lines: Critical Sets and Maximum Number of Solutions," IEEE Conf. on Computer Vision and Pattern Recognition, pp. 254-260, New York, June, 1993.

[23] A. R. Pose and 1). G. Lowe, "Vista: A Software Environment for Computer Vision Research," Int. Conf. on Computer Vision and Pattern Recognition, 1994.

[24] G. L. Scott, *Local and Global Interpretation of Mowing Images*, Morgan Kaufman Publishers, Inc., Los Altos, CA.

" *[25]* A. Zakhor, "Edge-Based 3-D Camera Motion Estimation with Application to Video Coding,>' IEEE Trans. on Image l'recessing, vol. 2, no. 4, pp. 481-498, 1993.

Figures

Fig, 1. An example of a calibrated overlay of both robot arm and ORU graphic models on the video image after the virtual reality calibration.

Fig. 2. An example of a preview/predictive display during the performance of the ORU extraction in the JPL/GSFC ORU changeout demonstration task.

Fig. 3. TELEGRIP's code and data architecture.

Fig. 4. TELEGRIPS'S Viewing model.

Fig. 5. a) A graphics model of the ORU and b) a video image.

Fig. 6. a) Multi-resolution edge-detected graphic and video images using the Canny operator, and b) correlation results.

Fig. 7. Detection of straight line segments after Canny and Lowe operators for a) the graphic model and b) video image.

Fig. 8. Video overlays a) before and b) after the edge-based feature matching.

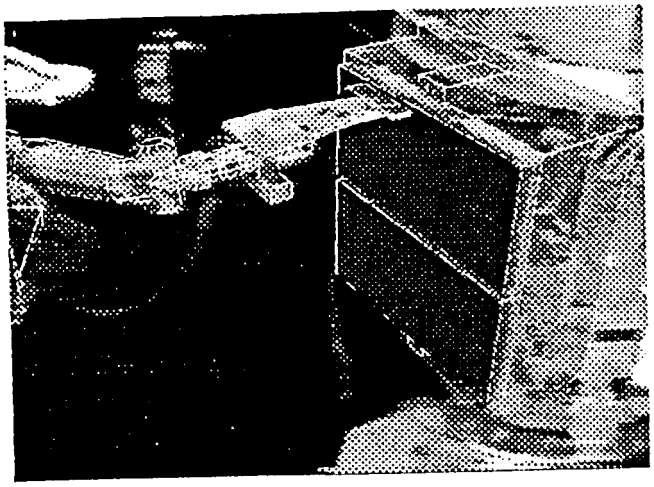

Figure 1: An example of a calibrated overlay of both robot arm and ORU graphic models on the video image after the virtual reality calibration

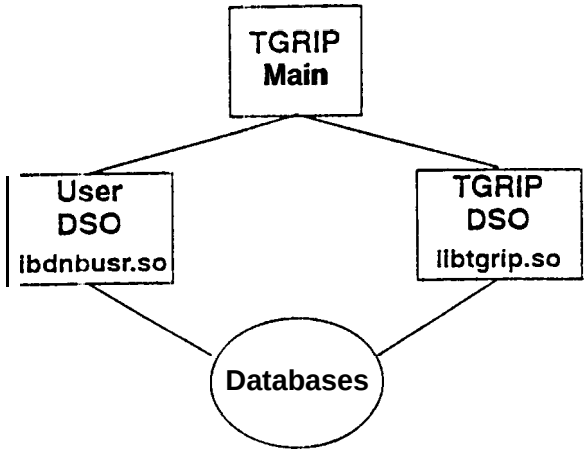

Figure 3. TELEGRIP's code and data architecture. TGRIP Main contains initialization code only. All other TELEGRIP functions are contained in libtgrip.so.

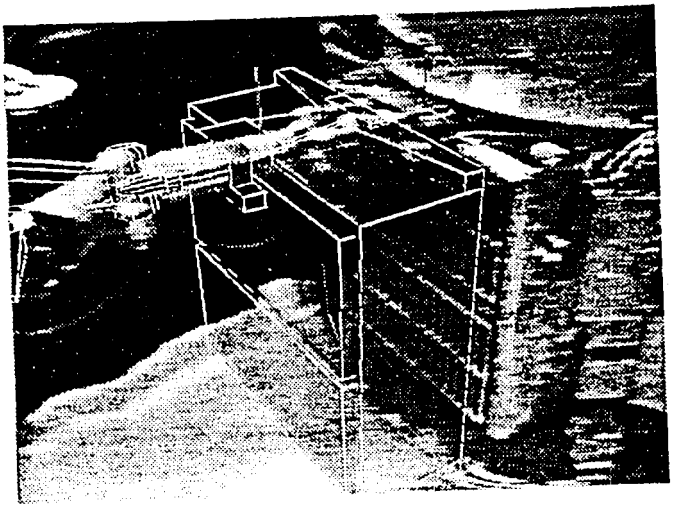

Figure 2: An example of a preview/predictive display during the performance of the ORU extraction in the JPL/GSFC ORU changeout demonstration task.

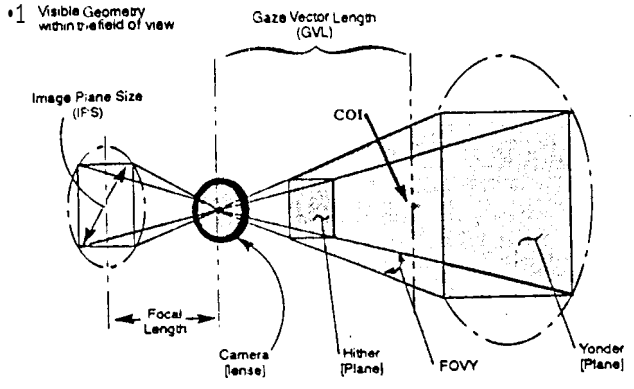

Figure 4: TELEGRIP's viewing model

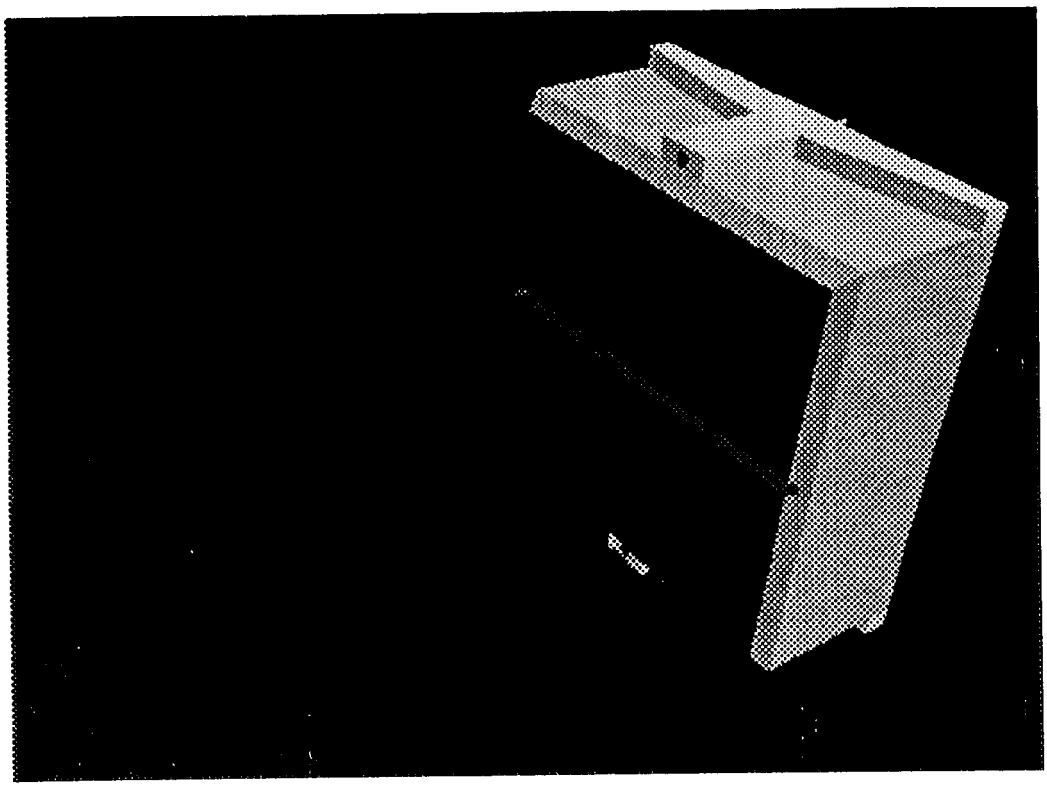

Fig. 5a

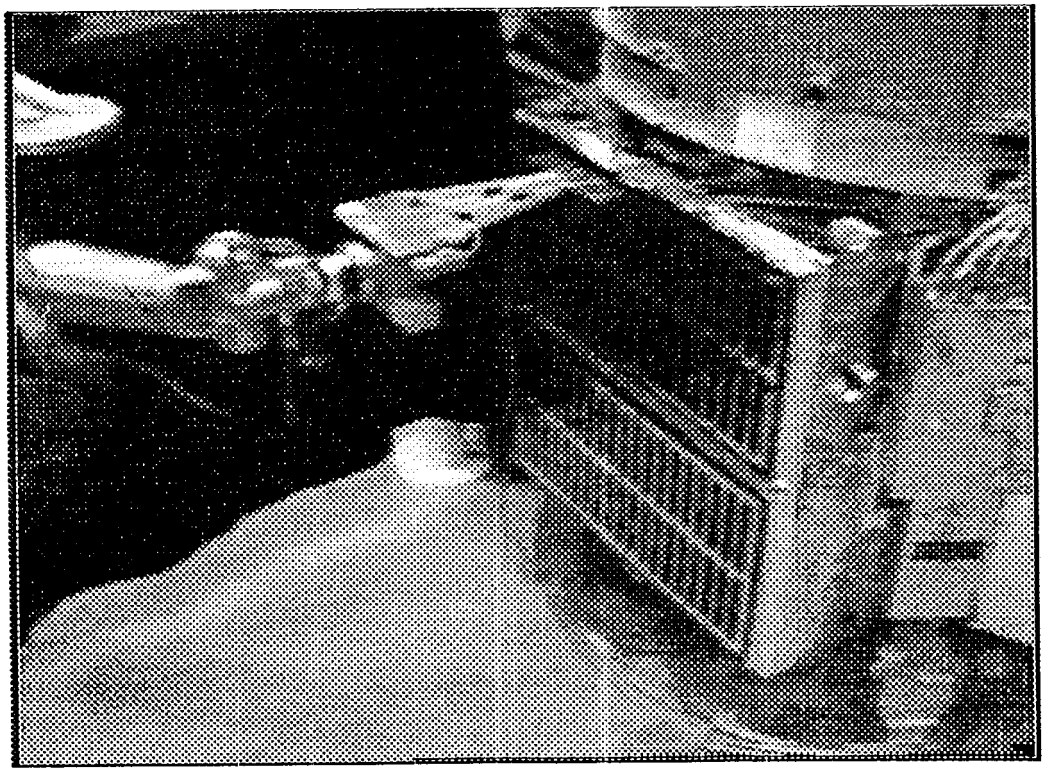

Fig 5b

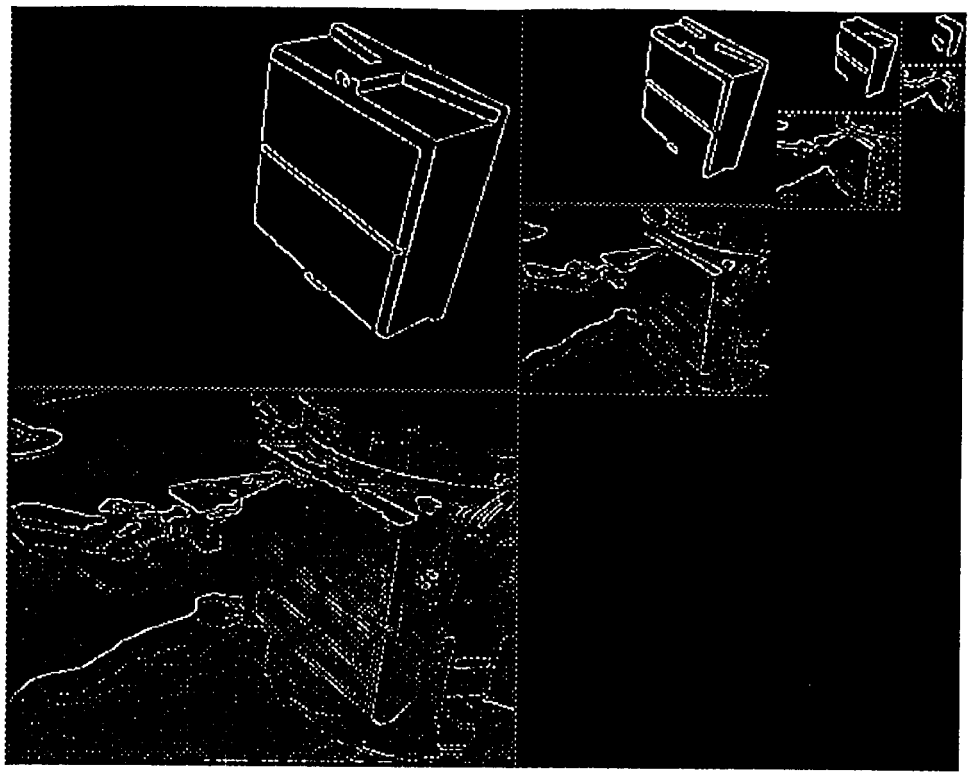

Fig. 6a

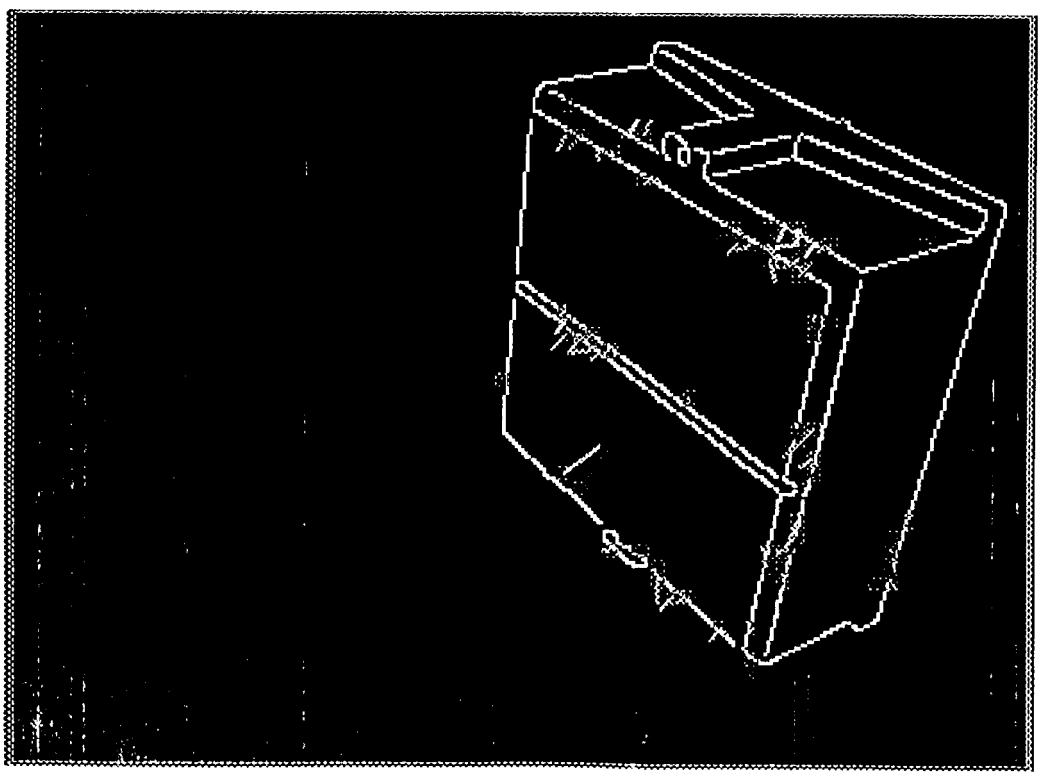

Fig. 6b

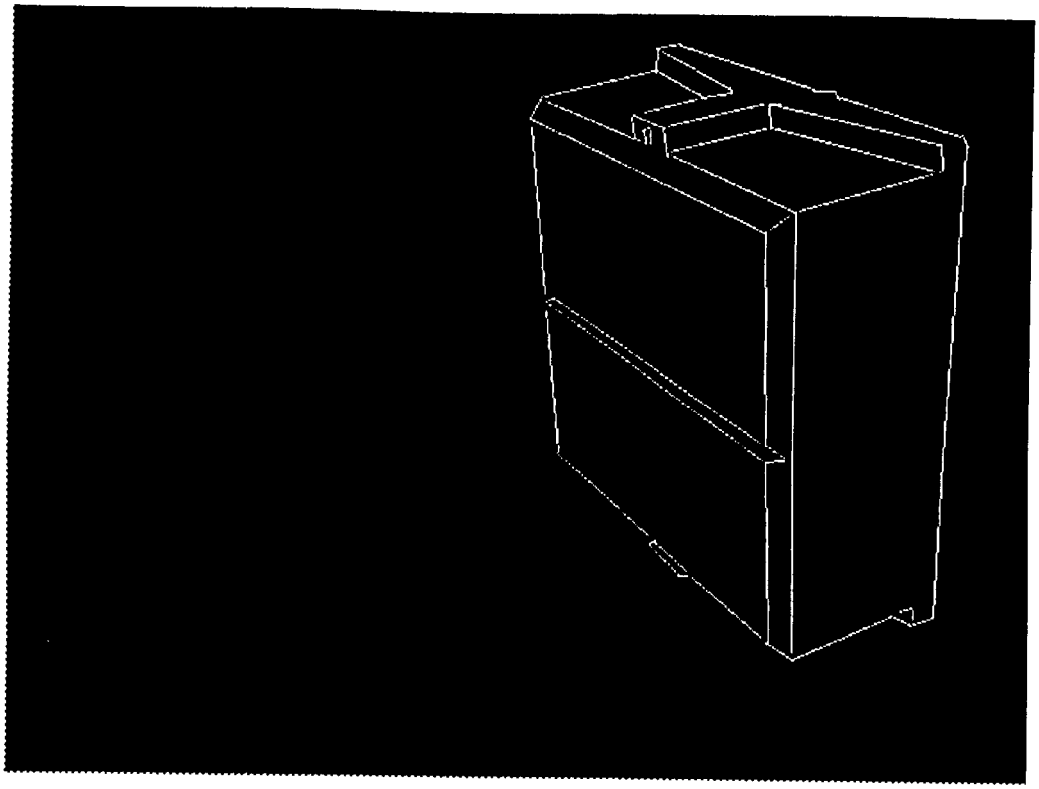

Fig. 7a

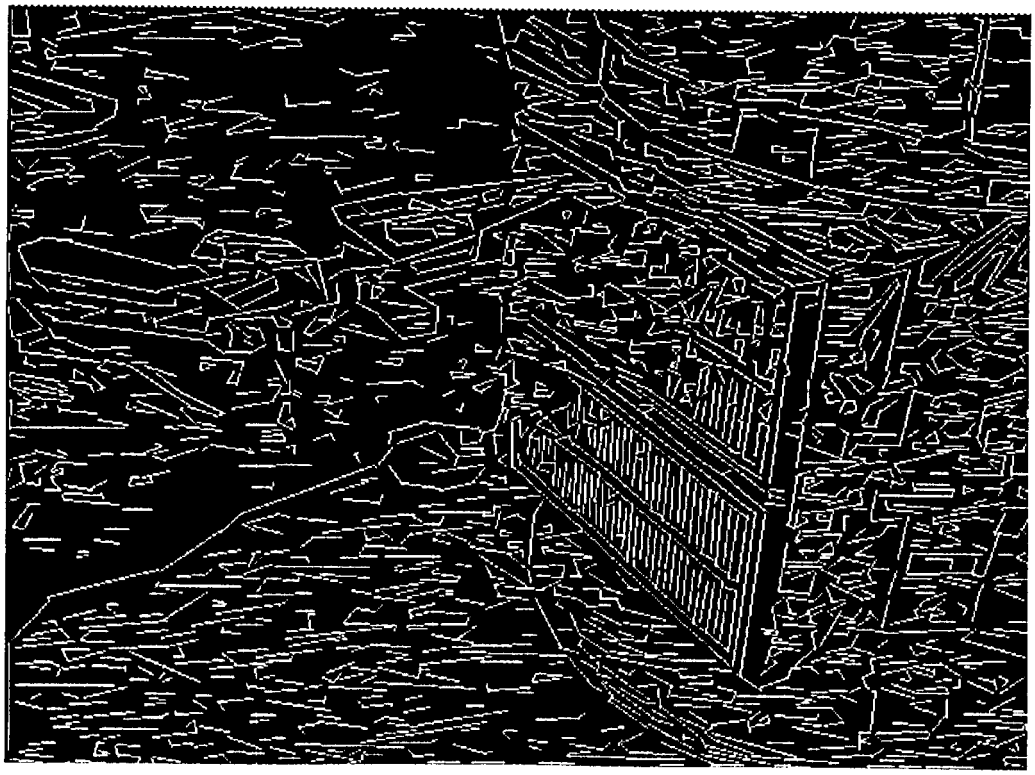

Fig. 7b

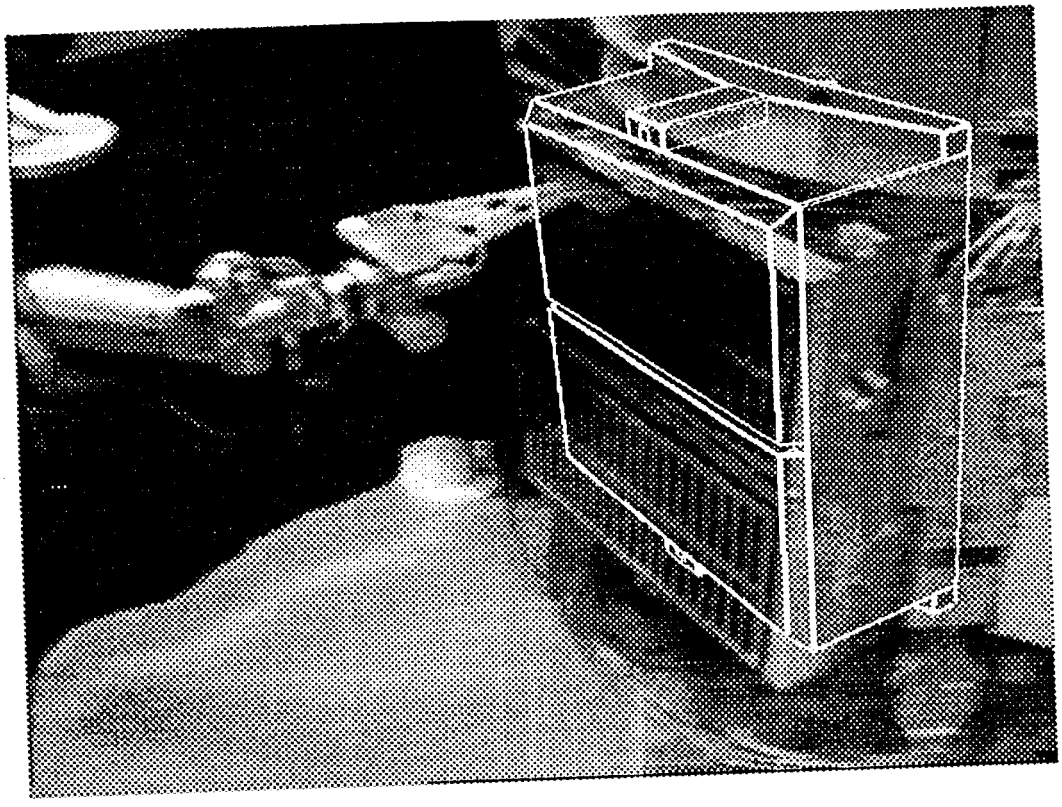

Fig. 8a

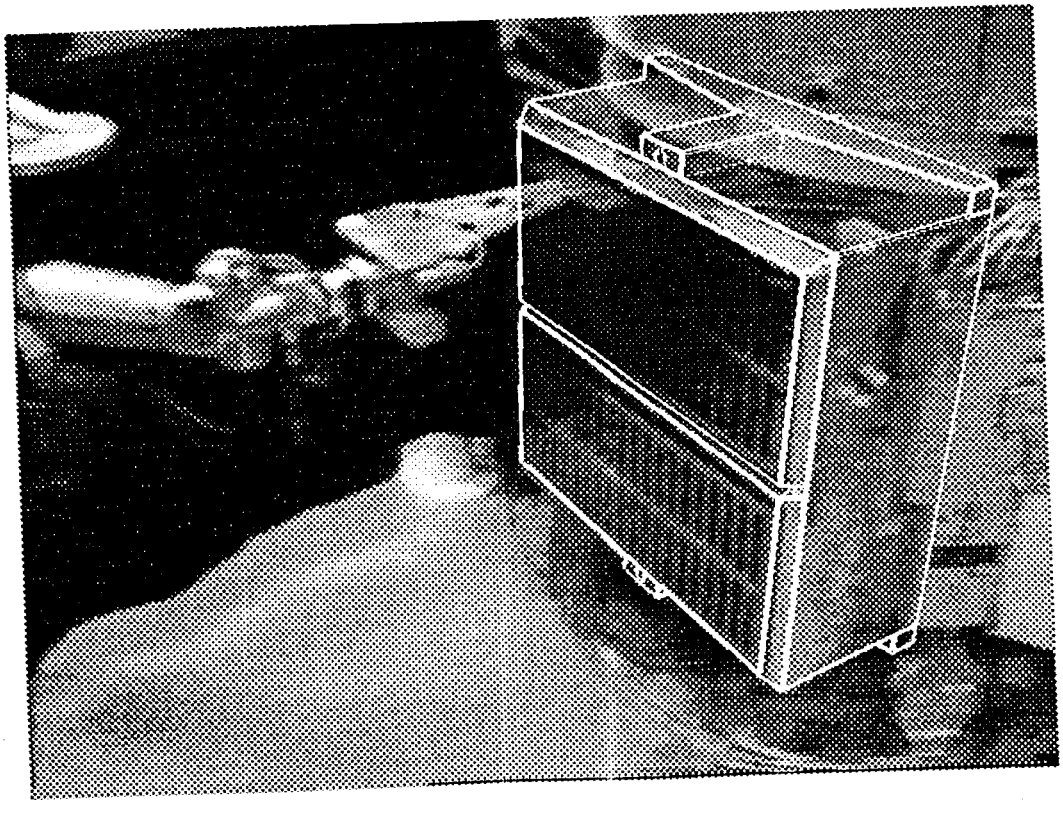

Fig. 8b